

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl. ⁶ G10L 11/00	(11) 공개번호 특2000-0032269 (43) 공개일자 2000년06월05일
(21) 출원번호 10-1998-0048673	
(22) 출원일자 1998년11월13일	
(71) 출원인 웰지전자 주식회사 구자홍 서울특별시 영등포구 여의도동 20번지	
(72) 발명자 이윤근 서울특별시 마포구 도화2동 우성아파트 6동 1405호 유수정	
(74) 대리인 경기도 성남시 분당구 구미동 무지개마을 주공아파트 1208-504 김용인, 심창섭	

설사청구 : 없음

(54) 음향 기기의 음성인식장치

요약

주변 잡음 속에서도 음향 기기를 제어할 수 있는 음향 기기의 음성인식장치에 관한 것으로, 음향 기기에 서 발생하는 노이즈를 적응 필터에 의해 제거하며, 음성이 입력되면 음성 구간 검출 알고리즘에 의해 음성 구간을 검출한다. 검출된 음성 구간에 대하여 음성을 분석하여 특징을 추출하고, 추출된 특징을 기준 패턴과 비교하거나 기준 모델의 발생 확률을 계산하여 입력된 음성을 인식 또는 거부하며, 추출된 특징값 및 인식 결과를 인식기에 적응하여 인식된 결과를 주변기기에 수행함으로써, TV, 오디오, 라디오 등의 음향 기기를 음성 인식에 의해 제어하여 편리성 및 부가가치를 증대시킬 수 있다.

대표도

도2

형세서

도면의 간단한 설명

도 1은 일반적인 음성 인식 시스템의 블록구성도
도 2는 본 발명에 따른 음향 기기의 음성인식장치의 블록구성도
도 3은 도 2의 적응 필터부를 보여주는 블록구성도

도면의 주요부분에 대한 부호의 설명

21 : 적응 필터부	22 : 음성 인식부
25 : 제 1 혼합기	26 : 적응 필터
27 : 제 2 혼합기	28 : 음성 구간 검출부
29 : 특징 추출부	30 : 매칭부
31 : 인식 결과 확인부	32 : 인식 결과 수행부

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 음성인식에 관한 것으로, 특히 주변 잡음 속에서도 음향 기기를 제어할 수 있는 음향 기기의 음성인식장치에 관한 것이다.

현재, 일상 생활에 많이 사용되는 전자 제품을 음성 명령으로 동작시키는 일이 여러 분야에서 시도되고 있다.

특히 TV, 오디오(audio) 등은 현대인의 생활에서 여가, 정보수집과 교육 등 중요한 부분을 차지하고 있다.

따라서, 작동의 편리성을 향상시키기 위해 리모콘이 등장하고 이로 인하여 앉은 자리에서 손으로 자신이 원하는 동작을 시킬 수 있게 되었다.

그러나, 이것 또한 다른 기기를 작동시키는 것이기 때문에 편리성의 한계가 있으므로 인간에게 가장 친숙한 음성을 인식할 수 있는 장치를 고안하게 되었다.

이러한 음성인식의 대략적인 동작원리는 다음과 같다.

먼저, 음성인식 알고리즘은 크게 음성 구간 검출 과정과 특징(feature) 추출 과정, 그리고 매칭(matching) 과정으로 나눌 수 있다.

즉, 도 1에 도시된 바와 같이 마이크(11)를 통해 음성신호가 입력되면 A/D 컨버터(12)에서 이를 디지털 신호로 변환한 후, 음성 구간 검출부(13)로 출력한다.

음성 구간 검출부(13)는 디지털 음성 신호를 짧은 구간의 신호(즉, 프레임)로 분할한 후, 각 프레임의 에너지와 영교차율(zero crossing rate) 그리고, 시간길이 정보를 이용하여 입력된 신호중에서 실제 발생된 음성 구간만을 검출한 후, 특징 추출부(14)로 출력한다.

특징 추출부(14)에서는 음성 구간에 해당하는 프레임의 특징을 추출하여 입력된 음성의 테스트 패턴을 만든 후, 매칭부(15)로 출력한다.

매칭부(15)에서는 테스트 패턴과 기준 데이터용 메모리(16)에 저장된 각 기준 패턴들과 각각 비교하여 테스트 패턴과 가장 유사한 특징을 갖는 기준 패턴을 인식된 음성으로 출력한다.

이때, 상기 기준 데이터용 메모리(16)에 음성 신호의 기준 패턴을 저장하는 경우도 마찬가지로 상기 특징 추출부(14)에서 음성 구간에 해당하는 프레임의 특징을 추출하여 기준 패턴을 만든 후, 기준 데이터용 메모리에 저장하는데, 이와 같은 동작을 인식하고자 하는 음성 신호에 대하여 반복 수행하여 기준 패턴을 기준 데이터용 메모리(16)에 데이터베이스화한다.

발명이 이루고자 하는 기술적 과제

그러나, 이와 같은 기존의 음성인식 장치는 주변 환경 잡음이 있을 경우 그 성능이 저하된다.

특히, TV, 오디오와 같이 음향 신호를 발생하는 기기에 마이크를 장착하여 음성인식을 할 경우에는 음향 기기에서 발생하는 음향의 크기가 매우 크므로 음성인식을 할 수가 없다.

따라서, 이러한 주변 음향의 존재에도 불구하고 음성인식이 가능한 장치의 개발이 필요하다.

본 발명은 이와 같은 문제를 해결하기 위한 것으로 음향 기기에서 발생하는 주변 음향에도 불구하고 음성인식이 가능한 음향 기기의 음성인식장치를 제공하는데 그 목적이 있다.

본 발명의 다른 목적은 음성인식 알고리즘을 마이크가 내장된 음향 기기에 적용하여 음성으로 음향 기기를 동작시키는데 있다.

발명의 구성 및 작용

본 발명에 따른 음향 기기의 음성인식장치의 특징은 음향 기기에서 발생한 노이즈 신호를 적어도 하나 이상의 마이크로부터 입력된 음성신호에 적응시켜서 상기 노이즈 신호가 제거된 음성신호를 출력하는 적응 필터부와, 적응 필터부의 음성신호를 인가받아 음성 구간 검출 알고리즘을 수행하여 음성 구간을 검출하고, 검출된 음성 신호를 분석하여 특징을 추출하며, 그 추출된 특징을 기준 패턴과 비교하거나 기준 모델의 발생 확률을 계산하여 입력된 음성을 인식 또는 거부하고, 상기 추출된 특징값 및 인식 결과를 인식기에 적용하여 인식된 결과를 주변기기에 수행하도록 하는 음성인식부로 구성되는데 있다.

본 발명의 다른 특징은 적응 필터부를 음향 기기에서 발생한 노이즈 신호와 마이크로부터 입력된 음성신호를 가산하여 노이즈가 혼합된 음성신호를 출력하는 제 1 혼합기와, 제 1 혼합기로부터 인가되는 노이즈가 혼합된 음성신호와 하기 적응 필터로부터 인가되는 필터링된 신호를 강산하여 노이즈가 제거된 음성신호를 출력하는 제 2 혼합기와, 인가되는 노이즈 신호를 필터링하여 출력하고 제 2 혼합기로부터 출력되는 노이즈가 제거된 음성신호를 피드백하여 필터의 계수를 조정하는 적응 필터로 구성하는데 있다.

본 발명의 또 다른 특징은 음성인식부에서 음성 구간 검출시, 적응 필터의 계수 변화 및 음향 기기에서 발생한 노이즈 신호와 마이크로부터 입력된 음성신호의 특징을 이용하는데 있다.

상기와 같은 특징을 갖는 본 발명에 따른 음향 기기의 음성인식장치를 첨부된 도면을 참조하여 설명하면 다음과 같다.

먼저, 본 발명의 개념은 음향 기기에서 발생하는 음향 신호(노이즈)를 기준 신호로 이용한 적응 필터(adaptive filter)를 구성하여 음향 기기에서 발생하는 노이즈(noise)를 제거함으로써 음향 기기에도 음성 인식을 가능하게 하는데 있다.

도 2는 본 발명에 따른 음향 기기의 음성인식장치를 보여주는 블록도로서, 도 2에 도시된 바와 같이 크게 적응 필터부(21)와 음성 인식부(22)로 나눌 수 있다.

적응 필터부(21)는 도 3에 도시된 바와 같이 음향 기기에서 발생한 노이즈와 마이크로부터 입력된 음성신호를 가산하여 노이즈가 혼합된 음성신호를 출력하는 제 1 혼합기(25)와, 제 1 혼합기(25)로부터 인가되는 노이즈가 혼합된 음성신호와 적응 필터(26)로부터 인가되는 필터링된 신호를 강산하여 노이즈가 제거된 음성신호를 출력하는 제 2 혼합기(27)와, 인가되는 노이즈를 필터링하여 출력하고 제 2 혼합기(27)로부터 출력되는 노이즈가 제거된 음성신호를 피드백하여 필터의 계수를 조정하는 적응 필터(26)로 구성된다.

여기서, 노이즈는 제어하려는 음향 기기에서 발생하는 음향 신호를 의미하는 것이고, 마이크로부터 입력된 음성신호는 외부의 노이즈 신호와 음향 기기에서 발생한 노이즈 신호가 포함된 음성신호를 의미한다.

또한, 마이크는 한 개 이상을 이용할 수 있으며, 음향 기기에 부착하거나 또는 음향 기기 외부의 다른 곳에 장착하여 음향 또는 음성 신호를 받아들인다.

그리고, 음성 인식부(22)는 적응 필터(26)의 계수 변화 및 마이크를 통해 입력되는 음성신호와 노이즈의 특징을 이용하여 마이크를 통해 인가된 신호로부터 음성 신호 구간을 추출하는 음성 구간 추출부(28)와, 검출된 음성 신호로부터 특성을 나타내는 파라미터를 추출해내는 특징 추출부(29)와, 추출된 파라미터를 기준 패턴과 비교하거나 기준 모델의 발생확률을 계산하여 입력된 음성을 인식하는 매칭부(30)와, 인식된 결과의 유사도 또는 발생 확률값이 일정 수준 이하이면 인식을 거부하는 인식 결과부(31)와, 인식된 결과를 수행하는 인식 결과 수행부(32)와, 추출된 특징값 및 인식 결과를 이용해 음성 인식 장치를 화자 또는 사용환경의 특성에 맞게 적응시키는 인식기 적응부(33)로 구성된다.

이와 같이 구성되는 본 발명의 동작을 도 2 및 도 3을 참조하여 설명하면 다음과 같다.

먼저, 음향 기기에서 음성 인식을 하기 위해서는 마이크에서 인가되는 음성신호로부터 음향 기기에서 발생하는 음향 신호(노이즈)를 제거해야 한다.

그러기 위해서는 음향 기기의 발생 신호와 마이크에서 인가되는 음성 신호의 차이를 최소화하도록 적응 필터 계수를 조정해야 한다.

즉, 음향 기기에서 발생한 노이즈와 마이크로부터 입력된 음성신호가 제 1 혼합기(25)에서 가산되어 노이즈가 혼합된 음성신호를 출력하고 적응 필터(26)는 노이즈 신호를 필터링한다.

이 적응 필터(26)의 출력 신호와 제 1 혼합기(25)의 출력 신호는 제 2 혼합기(27)로 인가되어 노이즈가 제거된 음성신호를 출력하게 된다.

그리고, 노이즈가 제거된 음성신호는 피드백(feedback)되어 적응 필터(26)로 다시 인가되어 적응 필터의 계수를 조정하게 된다.

이와 같이 적응 필터부(21)로부터 출력된 노이즈가 제거된 음성 신호는 음성 구간 검출부(28)로 인가되어 그 음성 신호 중에서 실제로 발생된 구간, 즉 음성 구간만을 검출한다.

여기서, 음성 구간은 적응 필터의 계수 변화 및 마이크로부터 입력되는 음성 신호와 음향 기기에서 발생하는 노이즈 신호의 특징을 이용하여 검출된다.

이들을 이용하는 이유는 적응 필터의 경우, 필터 계수는 기준 신호와 마이크로부터 입력되는 주신호가 현저하게 다를 경우 변화가 매우 심하게 나타난다.

즉, 음성이 입력되는 동안에 필터 계수 변화는 다른 구간과 현격한 차이를 보이게 된다.

그러므로 계수 변화량을 계속 모니터링하게 되면 변화량이 큰 부분을 음성 구간으로 추출할 수 있는 것이다.

여기서, 각 계수 변화량은 단구간 평균 계수 변화량을 측정한 것이다.

이러한 필터 계수 변화량과 더불어 기준 신호와 주신호의 스펙트럼 정보, 에너지 정보 등을 같이 이용하여 음성 구간을 추출한다.

이어, 특징 추출부(29)에서는 검출된 음성 구간 신호의 특징을 추출하게 되는데, 보통 스펙트럼 정보를 나타내는 파라미터를 이용한다.

그리고, 매칭부(30)는 추출된 파라미터를 기준 패턴과 비교하거나 기준 모델의 발생확률을 계산하여 입력된 음성을 인식한다.

즉, 입력된 음성과 가장 유사한 음성을 등록되어 있는 음성중에서 찾는데, 등록되어 있는 각 단어에 대한 유사도를 모두 계산한 뒤, 가장 유사한 등록 단어를 추출한다.

유사도를 측정하는 방법의 예로 DTW(Dynamic Time Warping) 알고리즘 또는 은닉 마코프 모델(Hidden Markov Model ; HMM) 확률값을 이용할 수 있다.

본 발명의 매칭부(30)에서는 음성인식 장치의 성능을 향상시키기 위하여 2단계 인식 방법을 이용할 수 있다.

2단계 인식은 키 워드(key word) 인식과정과 명령어 인식과정으로 나뉘는데, 키 워드 인식과정은 미리 정해진 한 가지 단어가 발생되었는가를 인식한다.

즉, 인식 대상 어휘는 1개이며 인식 결과 확인과정에 의해 '인식' 또는 '거부'를 결정하는데, 인식 단어를 최소화함으로써 인식율을 향상시킨다.

제어 대상 기기는 초기 상태로써 키 워드 인식 모드로 동작하고, '인식'으로 결정되면 제어하려는 기기를 명령어 인식 모드로 설정하고 기기에서 발생하는 노이즈를 일정 레벨(level) 이하로 제한한다.

그리고, 명령어 인식 과정은 기기의 명령어를 인식하는데, 인식 결과를 수행하고 기기를 본래의 상태인 키 워드 인식 모드로 설정한다.

이어, 인식 결과 확인부(31)는 가장 유사한 등록 단어와 입력 음성의 유사도가 일정 레벨에 의해 거부될 경우, 등록되어 있지 않은 음성이 입력된 것으로 보고 다시 음성을 받아들인다.

이때, 단순한 유사도를 측정하여 이용할 수도 있으며 여러 가지 형태의 필터 모델을 이용할 수도 있다.

이는 키 워드 인식 과정 및 명령어 인식 과정에 모두 적용된다.

그리고, 인식 결과 수행부(32)는 인식 과정이 성공하면 그 결과에 따라 명령을 수행한다.

한편, 인식기 적응부(33)는 추출된 특징값 및 인식 결과를 이용해 음성 인식 장치를 화자 또는 사용환경의 특성에 맞게 적응시킨다.

이 과정은 기기의 실제 사용 이전의 훈련과정으로써 이루어지거나 또는 실제 사용하는 도중에 이루어질 수 있다.

발명의 효과

본 발명에 따른 음향 기기의 음성인식장치는 다음과 같은 효과가 있다.

첫째, TV, 오디오, 라디오 등의 음향 기기를 음성 인식에 의해 제어함으로써, 그의 편리성 및 부가가치를 증대시킬 수 있다.

둘째, 음향 기기에 음성 인식 기능을 적용함에 있어 가장 큰 문제였던 노이즈 문제를 해결함으로써, 음향 기기에서 발생하는 노이즈에도 불구하고 음성 인식을 가능케 할 수 있다.

(57) 청구의 범위

청구항 1

음향 기기에서 발생한 노이즈 신호를 적어도 하나 이상의 마이크로부터 입력된 음성신호에 적용시켜서 상기 노이즈 신호가 제거된 음성신호를 출력하는 적용 필터부와,

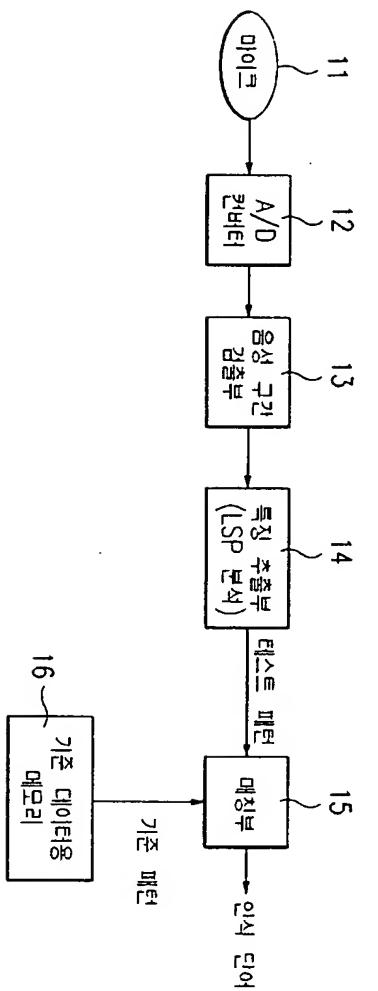
상기 적용 필터부의 음성신호를 인가받아 음성 구간 검출 알고리즘을 수행하여 음성 구간을 검출하고, 검출된 음성 신호를 분석하여 특징을 추출하며, 그 추출된 특징을 기준 패턴과 비교하거나 기준 모델의 발생 확률을 계산하여 입력된 음성을 인식 또는 거부하고, 상기 추출된 특징값 및 인식 결과를 인식기에 적용하여 인식된 결과를 주변기기에 수행하도록 하는 음성인식부로 구성되는 것을 특징으로 하는 음향 기기의 음성인식장치.

청구항 2

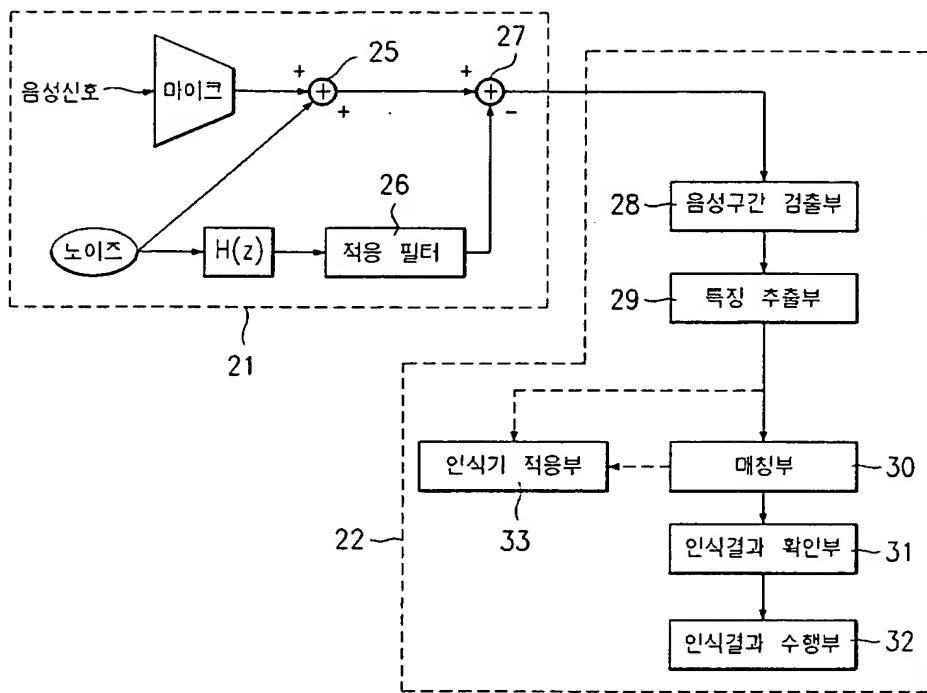
제 1 항에 있어서, 상기 적어도 하나 이상의 마이크로부터 입력된 음성신호는 외부의 노이즈 신호와 음향 기기에서 발생한 노이즈 신호가 포함된 음성신호인 것을 특징으로 하는 음향 기기의 음성인식장치.

도면

도면 1



도면2



도면3

